# APPENDIX D

## Current Efforts Work Group Report

The National Education Longitudinal Study of 1988 (NELS:88) is a comprehensive national survey administered by the National Center for Education Statistics in the U. S. Department of Education. The NELS:88 survey asks students if they participated in 4-H programs along with a menu of about 7-8 other school club and leadership opportunities. This would allow a secondary analysis of NELS:88 data comparing outcomes of students who indicated that they had participated in 4-H with those who had not. The NELS:88 database has approximately 5,000 variables and 26,000 students so this could be a comprehensive and nationally-representative study.

The base year of NELS:88 represents the first stage of a major longitudinal effort designed to provide trend data about critical transitions experienced by students as they leave elementary school and progress through high school and into postsecondary institutions or the work force. The 1988 eighth-grade cohort is being followed at two-year intervals. Policy-relevant data about educational processes and outcomes will be collected over time, especially as it pertains to student learning, early and late predictors of dropping out, and school effects on students' access to programs and equal opportunity to learn.

The first follow-up in 1990 constitutes the first opportunity for longitudinal measurements from the 1988 baseline. It also provides a comparison point to high school sophomores-ten years before, as studied in HS&B. The study captures the population of early dropouts (those who leave school prior to the end of tenth grade), while monitoring the transition of the student population into secondary schooling.

The second follow-up took place early in 1992, when most sample members were in the second term of their senior year. The second follow-up provides a culminating measurement of learning in the course of secondary school, and also collects information that will facilitate investigation of the transition into the labor force and postsecondary education after high school. Because the NELS:88 sample was freshened to represent the high school class of 1992, trend comparisons can be made to the high school classes of 1972 and 1980 that were studied in NLS-72 and HS&B. The NELS:88 second follow-up returned to students who were identified as dropouts in 1990,and identified and surveyed additional students who had left school since the prior wave.

The third follow-up took place in 1994, when most sample members had completed high school. The primary goals of the 1994 round were:
1) to provide data for trend comparisons with NLS-72 and HS&B;
2) to address issues of employment and post-secondary access and choice; and
3) to ascertain how many dropouts have returned to school and by what route. A fourth follow-up is tentatively scheduled for 2000.

# APPENDIX E

## *Methods/Pilot Work Group Report*
*INCLUDING PILOT RESULTS*

### *I. Pre-Pilot Phase (Summer and Fall 1998)*

In July of 1998, members of the Methods Working Group met in Ft. Collins, Colorado, to draft the sampling design and survey instruments. The team members referred to the Critical Elements and Conceptual Outcomes documents from the first two working groups. We also considered other 4-H and extension surveys (e.g., CYFAR reports, NY Members Only Club Survey, Iowa Survey, Kansas Survey, New Mexico and others) as well as non-4-H surveys (e.g., Boys & Girls Club, Big Brothers/Big Sisters, Girl Scouts, and others). Based on ideas from all these sources, many sample items were written for each critical element category and the items were ranked with the top 5-10 in each category used to draft the pilot instruments.

Arizona and Missouri helped develop the facilitated process and pre-piloted the process and the surveys with groups representing the four ES-237 programming units. In addition to the feedback from the Arizona and Missouri youth groups, NAE4-HA '98 workshop participants provided detailed suggestions for improvement of the surveys. Following the changes made from these sources the Pilot Phase was started in December, 1998.

### *II. Pilot Phase (Winter 1998-1999)*

***The purpose of the "Pilot Study" was to provide enough data to fine tune the survey instrument. A secondary purpose of the pilot phase was to try out the sampling process and to learn enough about the assessment effort to write a detailed "Instruction Packet" to provide to states involved in the national data collection effort.***

## Pilot Sampling Design

Selection Process:

- Randomly selected 2 states from each of the four regions
- Two counties were randomly selected from participating states for each of 3 program units (clubs, special interest and school enrichment) ensuring both rural and urban participation (See reference to Beale Index below*). After-school selections were determined whether a state had that type of program or not. The same instrument was used for all programming units.
- Each county was asked to prepare a list of all program groups within the selected unit in their respective counties from which two or more groups were selected by the assessment team to represent rural and urban audiences.
- Each county selected was asked to assess participants in one programming unit (i.e., 4-H Clubs)

The initial plan was to survey 2,000 youth participants as well as adult volunteers working with the groups. We sent out 1300 surveys to participating states and, chiefly due to the timing constraints of the pilot process, only about 480 youth and 190 adult surveys were returned. An additional number of completed surveys could not be used due to lack of written consent.

*Reference to Beale Index from USDA: Beale C. L. and K. M. Johnson, 1998. "The Identification of Recreational Counties in Nonmetropolitan Areas of the U.S.A." *Population Research and Policy Review.* 17: 37-53.

## Sample characteristics from pilot study

Valid surveys were received from every participating state (Rhode Island, Pennsylvania, Minnesota, North Dakota, Georgia, Texas, Oregon and Idaho). The youth ages ranged from 5-19 years with an average age of 9 and most of the youth in the 6th grade. The adult ages ranged from 21-57 years with an average age of 40. Seventy-three percent of the adults were parents and 53% were adult volunteers with 4-H (note that these are not mutually exclusive categories). Thirty-six percent of the adults were 4-Hers in their youth. The majority of both youth and adult were white (86% of youth and 92% of adults), although all ethnic categories were represented.

The samples programming sources were skewed because rural clubs were best able to respond to the tight time frame of the pilot study.

However, the experience/involvement of the youth in 4-H was highly varied. Their years of experience in 4-H ran the gamut from less than 6 months to more than 6 years. They were involved in clubs (73% reported participation in multi-project club) and other 4-H programs (e.g., 42% had done after-school programs, 32% school enrichment, 35% special interest, 31% 4-H overnight camp). Given that we needed numbers to look at the survey characteristics, ample responses were received. Difficulties experienced helped to craft the Instruction Packets for future use, the pilot phase experience was successful.

## *Survey Instrument Development*

Data from the pilot surveys was used to develop the final instrument. Two procedures—Factor Analysis and Cronbach's alpha—were used to select the final items for the attitude scales in the final version of the survey instrument. Factor analysis results were of minimal value in that most items grouped into the first factor indicating that the various aspects of the 4-H program are not mutually exclusive components. Cronbach's alpha is based on the acceptable range for all scales. The number of items in some scales based on the alpha scores were reduced. Alpha and mean scale scores from the pilot data for are as follows:

| *Scale Name* | *Alpha Value* | *Mean* | *Standard Deviation* | *N Valid* |
|---|---|---|---|---|
| *Adults in 4-H* | *.85* | *3.2* | *0.4* | *415* |
| *Feelings about 4-H* | *.84* | *3.3* | *0.4* | *433* |
| *Learning in 4-H* | *.77* | *3.1* | *0.5* | *442* |
| *Helping others* | *.87* | *3.2* | *0.5* | *443* |
| *Belonging in 4-H.* | *. 89* | *3.3* | *0.4* | *429* |
| *Planning and decision making in 4-H* | *.90* | *3.2* | *0.4* | *454* |

## *Data Analysis*

*(Some suggestions for those states interested in doing their own analysis)*

Data analysis can take several directions depending on the purpose of the survey and the level of complexity of the report. Once the data has been entered into the computer the first level of analysis is looking at a frequency distribution of all the variables in the survey. This will provide a general idea of the composition of the sample and their perceptions of the critical element items. For some this level of analysis may be sufficient. It provides a general overview of the data but does not give any information as to how perceptions vary depending on the number of years in 4-H or any of a number of background variables. Some may wish to move to the next level of analysis—a bivariate analysis. The most common bivariate analysis is a crosstabulation analysis or contingency table analysis. This allows one to evaluate the relationship between an independent and dependent variable. An example would be to look at whether perceptions (dependent variable) toward adults in 4-H is correlated with the number of years the person has been in 4-H (independent variable). This can be done with individual items in the "Adults in 4-H" section of the survey. However, a more preferred approach is to develop a scale from all the items in each section and use it as the dependent variable in the analysis. This can be done by totaling the scores for all items and dividing by the number of completed items to obtain an average for each section.  SPSS can be programmed to develop these scales. In most cases it is wise to group both the independent and dependent variables into 3-4 categories with roughly the same numbers in each category. This will assist in determining trends and minimizing the likelihood of drawing conclusions based on small cell sizes.

A bivariate analysis should never be the final approach to analysis of the data as there are too many opportunities for misinterpretation of the results. After all there are few dependent variables in the social sciences that can be explained by a single independent variable and too many independent variables that are correlated with other independent variables. Thus, the use of a multivariate procedure is advisable. This allows one to determine the relative influence of each independent variable on the dependent variable of choice, while holding all other independent variables constant. Several procedures can be used. Multiple Regression is a common procedure. Log Linear analysis techniques are also appropriate. A seldom used but very appropriate technique is Multiple Classification Analysis. This variation of ANOVA is quite similar to multiple regression but does not require the stringent assumptions of MR and actually provides more information about the relationship of the independent variable and dependent variable than does MR. Requirements for MCA are an interval level dependent variable and independent variables of any level of measurement from nominal to interval. Independent variables should be grouped into 3-5 categories for easier interpretation of results. In older versions of SPSS (prior to version 8), MCA analysis was on a pull down menu under the Simple ANOVA program. In more recent versions one must type the command file (Syntax file) to run the program. The commands can be obtained from Barbara Schnabel at the SSRU at the University of Idaho (bschnabe@uidaho.edu). Help with interpretation of MCA can also be obtained from this source. See also description below.

As a general principle, the more information you want from your data, the more complex your statistical analysis must be A frequency distribution of all variables is a must and can provide a useful general picture of your data. However, multivariate analysis is necessary if you are interested in the various relationships among dependent and independent variables. Knowledge of the independent effect of length of time involved in 4-H on satisfaction with each of the critical elements was of interest. This kind of analysis can only be done using a multivariate technique.

## *MULTIPLE CLASSIFICATION ANALYSIS* *(An example)*

The usual analysis of variance table provides only the statistics necessary for significance testing. The fact that the effect of factor 'A' is significant merely indicates that the mean of at least one category of the factor 'A' is different from the grand mean, after appropriate adjustments are made. It is therefore important to examine the pattern of A's relationship to the criterion variable.

The Multiple Classification Analysis (MCA) table is a method of displaying the results of analysis of variance especially when there are no significant interaction effects. It is particularly useful when the factors examined are attribute variables that are not experimentally manipulated and therefore are correlated. Given two or more interrelated factors, it is valuable to know the net effect of each variable when the differences in the other factors are controlled for.

For example (refer to MCA table), suppose that the criterion variable is weekly wages of company employees and that factors are sex and race. The researcher is interested in the effects of these two factors because some discrimination is suspected. However, wages are also determined considerably by the level of education of the employee and the duration of employment. Therefore, the two variables (level of education and duration of employment) are introduced as covariates.

| Variable + Category | | Deviations from the Grand Mean Unadjusted | Adjusted for independents | Adjusted for independents & Covariates |
|---|---|---|---|---|
| Race | | | | |
| | White | +10 | +6 | +4 |
| | Non-white | - 40 | -24 | -16 |
| | (Eta and beta) | (.632) | (.384) | (.253) |
| Sex | | | | |
| | Male | +12 | +8 | +66 |
| | Female | - 18 | - 12 | - 9 |
| | (Eta and beta) | (.465) | (.310) | (.232) |
| Multiple R | | - - - | .648 | .866 |
| R2 | | - - - | .420 | .750 |

*MCA Table:*
*WAGES (Weekly Wages expressed as Deviation from the grand mean*
*BY SEX (Sex of the Employee), RACE (Race of the Employee)*
*WITH EDUC (Educational Attainment), and LENGTH (Length Employed by Company)*
*Grand Mean = 100 (in dollars)*

The numbers in the first column are the deviations of each category mean from the grand mean. In calculating these values, do not adjust for other factors or for covariates. The numbers in the second column indicate the adjusted mean values for each category (again expressed as deviations from the grand mean) adjusting for the other factor. Note the changes in these values: The effect of each factor diminishes as we adjust for the other factor, which suggests that sex and race are related (in the context of employment). It shows that male employees tend to be white, while female employees tend to be nonwhite. As adjustment is made for the two covariates in addition to race and sex, the effects of sex and race are reduced still further. As the numbers in the final column suggest, there are still substantial degrees of discrimination between races and between sexes.

One important use of MCA scores is to examine the pattern of changes in the effects of a given variable as we introduce more variables as controls. For example, there was initially a $50 difference between whites and nonwhites (+10 for Whites and -40 for Nonwhites). Some of this difference is due to the confounding effects of sex and probably differences in the educational level of the two race categories. When the confounding effects of sex are controlled, there remains a $30 difference between races (+6 for Whites and -24 for Nonwhites), and when the differences in education and length of employment are further controlled, this difference reduces to $20 (+4 for Whites and -16 for Nonwhites).

Another descriptive statistic of interest is the partial beta. If a new variable for each factor by assigning the MCA scores to each category (any effect-proportional coding, that is, any linear transformation of the original MCA scores, will do), the resultant standardized partial-regression coefficient is partial beta. It is informative to compare the original eta (which is equivalent to a simple beta from the bivariate linear regression of the dependent variable on the factor) with the partial betas resulting from first controlling for the other factors and then, in addition, controlling for the covariates.

For example, for the factor Race in the above example, the betas decrease from .632 to .253 as we introduce other controls. Finally, the multiple R at the bottom of the table indicates the overall relationship between the criterion variable and the independent variables. R2 in the second column represents the proportion of variation in Wages explained by the additive effects of Race and Sex; R2 in the last column represents the proportion of variation in Wages explained by the additive effects of all factors and covariates.

(If there is strong interaction between factors, the MCA scores become meaningless. The user therefore is advised to check for the significance of interaction effects before examining the MCA tables.)

### *MCA using SPSS*

To run an MCA using SPSS version 8.0 or higher you will need to create a SYNTAX file. On the menu for to File-New-Syntax. At the blank screen for a New Syntax file, type in:

```
ANOVA Y BY A (1,2) B (1,4)
 /METHOD=EXP
 /STATISTICS=MCA
 /MAXORDERS=NONE.
```

Y is the dependent variable (above example the dependent variable was Wages), A is the first independent variable with values ranging from 1 to 2 (above example this would be Race), B is the second independent variable with values ranging from 1 to 4 (example could be Education 1=High School, 2=Vocational Certificate, 3=College Degree, and 4=Graduate Degree). It is usually a good idea to limit the number of independent variables to about six and to limit value ranges either by specifying only part of the values (example left out 5=no answer) or by transforming (recoding) the variable.

Maxorders limits the number of interactions. This command is optional, but if it is not included and there are more than five independent variables with more than three values each, then the program might overload computer memory capacity.

***Reference:***

Statistical Package for the Social Sciences, Second Edition. Nie, Norman H.,. Hadlai C. Hull, Jean G. Jenkins, Karin Steinbrenner, and Dale H. Bent. 1975. McGraw-Hill Book Company, New York, NY.